Improved Low Resolution Heterogeneous Face Recognition using Re-ranking

Sivaram Prasad Mudunuri, Shashanka Venkataramanan, and Soma Biswas

Department of Electrical Engineering
Indian Institute of Science, Bangalore.
{sivaramm,shashankv,somabiswas}@iisc.ac.in

Abstract. Recently, near-infrared to visible light facial image matching is gaining popularity, especially for low-light and night-time surveillance scenarios. Unlike most of the work in literature, we assume that the near-infrared probe images have low-resolution in addition to uncontrolled pose and expression, which is due to the large distance of the person from the camera. To address this very challenging problem, we propose a re-ranking strategy which takes into account the relation of both the probe and gallery with a set of reference images. This can be used as an add-on to any existing algorithm. We apply it with one recent dictionary learning algorithm which uses alignment of orthogonal dictionaries. We also create a benchmark for this task by evaluating some of the recent algorithms for this experimental protocol. Extensive experiments are conducted on a modified version of the CASIA NIR VIS 2.0 database to show the effectiveness of the proposed re-ranking approach.

Keywords: Heterogeneous face recognition, dictionary learning and re-ranking.

1 Introduction

The task of recognizing low resolution facial images captured in uncontrolled settings has become a challenging area of research in the field of computer vision due to the increasing usage of surveillance cameras [1],[2]. It has been observed that the performance of most of the current algorithms degrades if the images captured are of low resolution and also contain varying poses and poor illumination conditions which replicates real world scenarios [3],[4],[5],[6].

To handle illumination variations at low-light or night-time conditions, researchers have started using near-infrared (NIR) facial images. But the gallery may consist of high resolution (HR), controlled visible images captured during enrolment. So our goal is to develop a heterogeneous face recognition algorithm which can match low resolution (LR) NIR face images captured under uncontrolled pose and expression with the HR visible face images captured with frontal pose and good illumination conditions (Fig. 1). This problem is more challenging since the probe and gallery faces differ in resolution, pose and illumination together with spectral variations. Addressing all these variations together has not been well studied and needs to be addressed with special attention since

Fig. 1. Low resolution heterogeneous face recognition: Matching HR VIS frontal faces (first row) with LR NIR non frontal faces (second row). Original images are taken from CASIA NIR VIS 2.0 database [8].

these replicate the practical scenario [7]. The proposed re-ranking module can be used as an add-on for any other algorithm also.

Recently, several approaches have been proposed for HR heterogeneous face recognition in the direction of dictionary learning framework. We build our re-ranking framework on the success of [9] to attempt this problem. Given a LR NIR probe and the HR visible gallery images, we select the k-nearest neighbors since even if the correct match is not at rank-1, it will likely be in the top few matches. We analyze the relative position of the probe and all the gallery in this set with respect to a set of reference images to re-rank these k-nearest neighbours. The assumption behind the re-ranking is that for the correct probe gallery pair, the nearest neighbors of both of them should match. The main contributions of our proposed work and differences with [9] are as follows:

- We provide a benchmark for the problem of LR heterogeneous face recognition and evaluate some of the recent approaches in literature for this task.
- We propose a re-ranking module that is evaluated as an add-on feature to an existing approach.
- The proposed re-ranking can potentially be used by other approaches to boost their performance.
- Extensive experiments conducted on a modified version of CASIA NIR VIS 2.0 database [8] show the effectiveness of the proposed approach.

2 Related Work

In this section, we provide some pointers regarding the recent approaches. A fusion based algorithm that uses restricted boltzman machine and stacked denoising auto encoders to address the problem of matching LR NIR and HR VIS is proposed in [7]. A computationally efficient correlation analysis involving discriminant correlation analysis to compute the projection matrices which maximizes the correlation is described in [5]. A joint face hallucination and recognition approach based on sparse representation is described in [10]. Moutafis *et al.* [11] propose a framework to match LR and HR

faces by jointly learning two semi-coupled bases to exploit the optimal representations. Li *et al.* [12] uses local binary pattern for low resolution face recognition by combining multi-scale blocking center symmetric local binary pattern based on Gaussian pyramid. Mudunuri *et al.* [2] propose an approach that learns multi-dimensional scaling to construct a common transformation matrix to simultaneously transform both the probe and gallery faces into a common discriminative space. A template based face recognition which efficiently fuses discriminative information of deep features is proposed in [13]. An inter-session variability modelling approach using Gaussian mixture model to handle cross-modal faces is described in [14]. Peng *et al.* uses Markov models to represent heterogeneous image patches which considers the spatial compatibility between neighboring patches. Xing *et al.* [15] describes a generalized bipartite graph to discretely approximate the manifold structure of face sets with different resolution.

A useful study on the effect of low resolution with deep learning techniques primarily involving pre-training by super resolution, domain adaptation and regression techniques is presented in [16]. A complex deep model to learn the relationship between cross-modal images is described in [17]. A single deep convolutional neural network architecture to map both NIR and VIS images to a compact Euclidian space is given in [18]. A neural network based architecture that captures the non-linear relationship between the two modalities of faces is proposed in [19]. A triplet loss based convolutional neural network framework is proposed in [20]. CNN based metric learning strategies to reduce discrepancies between the different modalities is proposed in [21].

3 Proposed Method

In this section, the details of the proposed approach is presented which addresses the problem of recognizing low resolution NIR faces captured under uncontrolled pose and expressions with HR visible (VIS) faces captured under frontal pose and good illuminations. The proposed re-ranking algorithm can be used as an add-on to any other algorithm to improve their matching accuracy for this problem. In this work, we use a recent work [9] as the baseline approach with certain modifications in which the authors addressed VIS-NIR face matching where NIR images are also of high resolution. For completion, first we will briefly describe the modified baseline approach [9] and then provide details of the re-ranking algorithm.

3.1 Baseline Approach

The baseline approach [9] has a training and a testing stage. During training, given images from both domains (NIR and VIS), we compute two orthogonal dictionaries for each of the domains respectively. Note that unlike [9], in this case, the NIR images are of low-resolution. Let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_1}\} \in \mathcal{R}^{d \times N_1}$ and $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_2}\} \in \mathcal{R}^{d \times N_2}$ be the data in the two domains, where d is the feature dimension of each sample and N_1, N_2 are the number of samples available for each of the respective domains. The main advantage of learning dictionaries separately for the two domains instead of in a coupled manner [22],[23] is that we can avoid the requirement of paired training data. The other advantages of learning orthogonal dictionaries over the over-complete

dictionaries are: 1) less computational complexity in computing the dictionaries and 2) redundancy of dictionary atoms is less.

The algorithm learns an orthogonal dictionary [24] $\bar{\mathbf{D}}_x = [\mathbf{J}_x \ \mathbf{D}_x]$ from the given data $\mathbf{X}$ such that $\bar{\mathbf{D}}_x^T \ \bar{\mathbf{D}}_x = \mathbf{I}_{d\times d}$. The optimization function is formulated as follows:

$$\begin{aligned} \min_{\mathbf{D}_x, \mathbf{\Lambda}} \| \mathbf{X} - [\mathbf{J}_x, \mathbf{D}_x] \mathbf{\Lambda} \|_2^2 + \alpha \| \mathbf{\Lambda} \|_0^2 \\ \text{subject to } \mathbf{D}_x^T \mathbf{D}_x = \mathbf{I}_{m\times m}, \mathbf{J}_x^T \mathbf{D}_x = \mathbf{0}. \end{aligned} \tag{1}$$

where d is the length of input feature vector and m is the number of atoms in the dictionary $\mathbf{D}_x$. The dictionary $\bar{\mathbf{D}}_x$ has two sub dictionaries: $\mathbf{D}_x$ is the orthogonal dictionary which has to be learnt from the given training data samples and $\mathbf{J}_x$ which can be used to control the required number of orthogonal dictionary atoms. The matrix $\mathbf{J}_x$ is set to a null matrix in our framework so that $\bar{\mathbf{D}}_x = \mathbf{D}_x$ and hence the dictionary atoms can span the entire space of d-dimensional vectors since they are orthogonal to each other (as per the constraint in equation (1)). The dictionary $\mathbf{D}_x$ is initialized to DCT matrix of size $d \times d$. During the i^{th} iteration, we have $\mathbf{D}_x^i$ and the task is to find out $\mathbf{\Lambda}_x^i$ and is formulated as below:

$$\begin{aligned} \hat{\mathbf{\Lambda}}_x^i = \arg\min_{\mathbf{\Lambda}_x^i} \left\| \mathbf{X} - \left[\mathbf{J}_x, \mathbf{D}_x^i\right] \mathbf{\Lambda}_x^i \right\|_2^2 + \alpha \left\| \mathbf{\Lambda}_x^i \right\|_0^2 \\ \text{subject to } \mathbf{D}_x^{i^T} \mathbf{D}_x^i = \mathbf{I}, \mathbf{J}_x^T \mathbf{D}_x^i = \mathbf{0}. \end{aligned} \tag{2}$$

The optimization problem in (2) has a unique solution [24] and is given by:

$$\hat{\mathbf{\Lambda}}_x^i = T_\gamma(\bar{\mathbf{D}}_x^{i^T} \mathbf{X}) \tag{3}$$

Here $T_\gamma(\mathbf{v})$ represents a hard threshold operation on the vector $\mathbf{v}$ and we experimentally tuned the parameter γ. The objective function to update the dictionary at i^{th} iteration can be formulated as:

$$\begin{aligned} \hat{\mathbf{D}}_x^i = \arg\min_{\mathbf{D}_x^i} \left\| \mathbf{X} - \left[\mathbf{J}_x, \mathbf{D}_x^i\right] \mathbf{\Lambda}_x^{i-1} \right\|_2^2 \\ \text{subject to } \mathbf{D}_x^{i^T} \mathbf{D}_x^i = \mathbf{I}, \mathbf{J}_x^T \mathbf{D}_x^i = \mathbf{0}. \end{aligned} \tag{4}$$

Since the two dictionaries are orthogonal and span the same/similar space in the two domains, we can assume that their exists a one-to-one correspondence of the dictionary atoms. Consider $\mathbf{D}_x$ and $\mathbf{D}_y$ be the the orthogonal dictionaries learnt from the two domains $\mathbf{X}$ and $\mathbf{Y}$. The one-to-one correspondence between the dictionary atoms is determined in a common space (here CCA space) and then the dictionaries are aligned similar to the subspace alignment approach [25]. A metric learning algorithm (here LSML [26]) is used to make the sparse coefficients discriminative. Please refer to [9] for more details of this baseline algorithm.

During the testing stage, given a low resolution NIR probe image, we need to identify the correct match from the database of gallery faces, which consists of HR visible images. In our work, the gallery and probe images differ in pose, illumination, resolution and spectral domain. We first compute the sparse coefficients from the respective aligned orthogonal dictionaries computed in the training stage for both the probe and gallery images. The sparse coefficients are transformed into the common space using the metric learnt using LSML approach.

3.2 Proposed Re-ranking algorithm

Given a probe, the above procedure gives a list of gallery images, where the first image has the highest similarity with the probe, followed by the second one and so on. Now, we describe the proposed re-ranking algorithm which works with this retrieved set of images for improving the matching accuracy. It is based on two assumptions, (1) Even though the correct match may not appear in the first rank, it may appear in the top k-nearest neighbours; (2) An image can be very well described by taking into account its relation with some reference set of images [2].

Many of the existing algorithms including the considered baseline algorithm uses the subject labels for retrieving the correct match. In this work, we use the additional information of the relation with the reference set to improve the matching performance. The idea is that though the subject is captured in either VIS or NIR mode, his/her relation w.r.t the selected reference set should not change significantly in either of the domains. Let $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_r}\} \in \mathcal{R}^d$ represent the features of N_r number of reference faces. If y_{te} is the probe face (sparse coefficient vector in the common space), then η= $[\eta_1, \eta_2, \ldots, \eta_{N_r}] \in \mathcal{R}^{N_r}$ is the vector which encodes the information (in our case distance) of how the probe face y_{te} is related to the set of reference faces. Each entry in the distance vector is computed as follows:

$$\eta_i = \| y_{te} - \mathbf{r}_i \|_2^2 \tag{5}$$

Similarly, we compute the vectors for the top k-nearest neighbors from the gallery faces for the given probe. Then the distances between the vectors of the probe and the k nearest galleries are computed. We take the weighted sum of previous distances (that are computed using transformed sparse coefficient vectors) and the new distances obtained with the computed vectors and then the first nearest neighbor is taken as the correct identity for the given probe. In our experiments, we used the training subjects itself as the reference subjects and we analyze the performance of our approach by varying the number of reference faces. In our experiments, we observe that 10 reference subjects give a reasonable boost in accuracy.

The importance of these distance vectors is illustrated in Fig. 2. In the figure, for a given probe (shown in left column), the top row shows the 5 most similar images in the gallery for the approach [9] and the bottom row shows the results using the proposed framework. If we take the second example in the figure, the correct identity (marked with red box) as per [9] appears at the fourth position. By employing the proposed re-ranking approach, the distances are improved so that the correct match comes to the first position.

4 Experimental Results

Extensive experiments are conducted on CASIA NIR-VIS 2.0 database [8] as this is the largest publicly available database for matching NIR faces with VIS faces. The dataset consists of VIS-NIR images from 725 subjects. The dataset was collected over the years in 4 sessions. There are 1-22 visual and 5-50 NIR images per subject. There are 10 training-testing splits recommended by the authors. The subjects in the training

Fig. 2. Illustration of retrieval performance of our proposed re-ranking framework and comparison with the method [9]. For each probe (left column), the top row shows the 5 most similar images in the gallery for the approach [9] and the bottom row shows the results using the proposed framework. The red boxes indicate the correct identity for the given probe.

and the corresponding testing set are non-overlapping, and the percentage of subjects in the training and testing set are both nearly 50%.

But in the database, both the VIS and NIR faces are high resolution images with size 128×128. So we converted the NIR faces into low resolution by down-sampling them to 20×20. They are then upsampled to the original size (128×128) so that the same features can be computed for all the images. We follow the same protocol as described in the database except that the NIR images used are of low resolution to mimic the real-world scenario. The rank-1 accuracy and the standard deviation of our approach over the 10 folds as per the protocol of the database is reported in Table. 1. In all the cases, VGG face features (pool5, fc6 and fc7 features) [27] are used as image features. To reduce the dimensionality, we applied PCA and took the first 2500 coefficients as the feature vector.

First, we evaluated several algorithms on this modified CASIA NIR-VIS 2.0 database to create the benchmark results. Specifically, we evaluated several correlation-based ap-

proaches, metric learning and dictionary learning approaches. We have directly taken the codes from respective authors. The results of these approaches and also of the baseline algorithm [9] is given in Table. 1 along with the results of the proposed re-ranking approach. We observe that the proposed re-ranking approach is able to improve upon the baseline approach and is superior to all the other compared approaches.

Table 1. Comparison of the performance (Rank-1 (%) $\pm$ Std) of the proposed approach on CASIA NIR-VIS 2.0 database [8] (NIR faces are of low resolution and the VIS faces are of high resolution) with state-of-the-art approaches.

Method	**Rank-1 (%) $\pm$ Std**
CBFD [28]	26.65 ± 0.35
GMA [29]	34.78 ± 0.89
CCA [30]	35.48 ± 1.46
Mean CCA [30]	35.93 ± 1.74
MvDA [31]	39.39 ± 1.86
Cluster CCA [30]	42.21 ± 1.55
Randomized Kernel CCA [32]	45.59 ± 1.13
Dictionary Alignment [9]	58.35 ± 1.37
Proposed Approach	$\mathbf{60.21 \pm 1.26}$

4.1 Analysis of the proposed approach

Now, we present some analysis done for the proposed approach.

Effect of different number of reference images: We evaluated with different number of reference images N_r which are used to compute the distance vectors and the variation of the Rank-1 recognition rate is given in Table 2. For this analysis, we took the third fold of the database protocol.

Table 2. Effect of increasing number of reference images N_r. Fold 3 is chosen to conduct the experiment.

N_r	10	20	30	50	100	200
Rank-1 (%)	57.20	58.37	58.71	58.74	58.92	59.38

Effect of applying super-resolution on LR NIR images: One way of matching LR

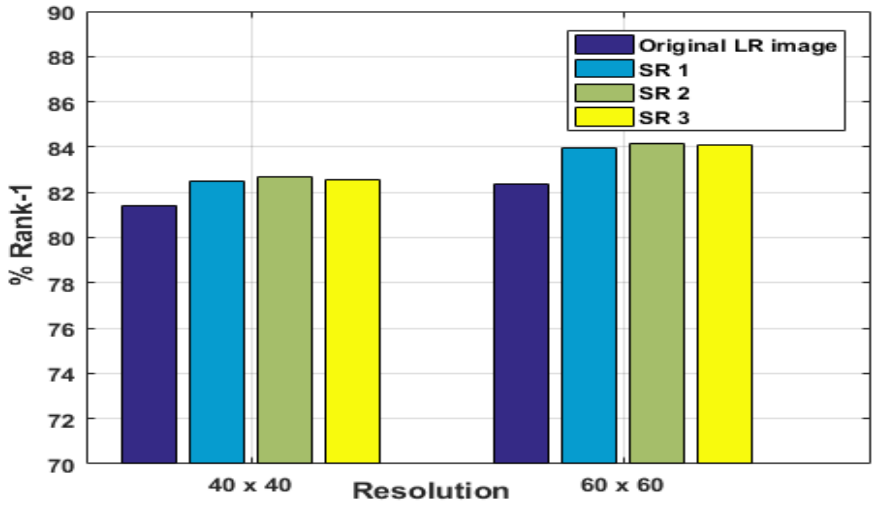

Fig. 3. Illustration of effect of applying super resolution techniques: SR1 [33], SR2 [34] and SR3 [35] on LR NIR faces before performing matching.

against HR faces is to first apply super resolution (SR) on LR images to improve their resolution and then perform the recognition. We experimented with three state-of-the-art super resolution approaches SR1 [33], SR2 [34] and SR3 [35] to make LR NIR images (of sizes 40×40 and 60×60) to HR and applied our approach to see the effectiveness. The resultant rank-1 recognition rates are shown in Fig. 3. We observe that the performance increases using SR techniques.

Effect of different probe resolutions: To analyze the effectiveness of the proposed approach across different probe resolutions, we evaluated the approach with different resolutions of probe faces. We observe from Table 3 that the performance remains stable till probe resolution of around 40×40 and then suddenly decreases drastically, which also demonstrates the importance of handling this problem.

Effect of LR visible images on the performance: We also analyze the effect of res-

Table 3. Performance of the proposed approach for different NIR probe resolutions.

Resolution	Rank-1 (%)
10×10	24.38
20×20	59.45
40×40	82.35
60×60	83.80
80×80	84.28

olution of the gallery VIS images. For this, we conducted experiments by varying the VIS face resolutions and the recognition accuracies are presented in Table 4. The NIR faces are fixed to the resolution 20×20. We observe that the accuracies are almost stable till gallery resolution of 20×20 and then suddenly drops.

Table 4. Performance of the proposed approach for different VIS gallery resolutions. Probe resolution is 20×20.

Resolution	**Rank-1 (%)**
10×10	24.59
20×20	51.36
40×40	58.51
60×60	58.16
80×80	59.03

5 Conclusion and Future Work

In this work, we proposed a re-ranking algorithm which can be used as an add-on to any existing algorithm for the problem of low resolution heterogeneous face recognition. The overall approach does not require one-to-one paired images for learning the dictionaries. Extensive experiments are conducted to show the usefulness of the proposed approach as well as the importance of the problem attempted in this work. Currently, there are no publicly available database which addresses this problem which has surveillance quality NIR images with large number of subjects. Collecting our own surveillance quality LR NIR database so as to advance research in this important field will be one of the future works.

Acknowledgment

This work is partly supported through a research grant from DeITY, India.

References

1. Zou, W.W., Yuen, P.C.: Very low resolution face recognition problem. IEEE Transactions on Image Processing, 21(1), pp.327–340 (2012).
2. Mudunuri, S.P., Biswas, S.: Low resolution face recognition across variations in pose and illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(5), pp.1034–1040 (2016).
3. Shekhar, S., Patel, V.M., Chellappa, R.: Synthesis-based robust low resolution face recognition. arXiv preprint arXiv:1707.02733 (2017).
4. Chu, Y., Ahmad, T., Bebis, G., Zhao, L.: Low-resolution face recognition with single sample per person. Signal Processing, 141, pp.144–157 (2017).

5. Haghighat, M., Abdel-Mottaleb, M.: Low Resolution Face Recognition in Surveillance Systems Using Discriminant Correlation Analysis. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 912–917 (2017).
6. Jiang, J., Hu, R., Wang, Z., Cai, Z.: Cdmma: Coupled discriminant multi-manifold analysis for matching low-resolution face images. Signal Processing, 124, pp.162–172 (2016).
7. Ghosh, S., Keshari, R., Singh, R., Vatsa, M.: Face identification from low resolution near-infrared images. In: IEEE International Conference on Image Processing, pp. 938–942 (2016).
8. Li, S., Yi, D., Lei, Z., Liao, S.: The casia nir-vis 2.0 face database. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 348–353 (2013).
9. Mudunuri, S.P., Biswas, S.: Dictionary Alignment for Low-Resolution and Heterogeneous Face Recognition. In: IEEE Winter Conference on Applications of Computer Vision, pp. 1115–1123 (2017).
10. Yang, M.C., Wei, C.P., Yeh, Y.R., Wang, Y.C.F.: Recognition at a long distance: Very low resolution face recognition and hallucination. In: International Conference on Biometrics, pp. 237–242 (2015).
11. Moutafis, P., Kakadiaris, I.A.: Semi-coupled basis and distance metric learning for cross-domain matching: Application to low-resolution face recognition. In: International Joint Conference on Biometrics, pp. 1–8 (2014).
12. Li, J., Chen, Z., Liu, C.: Low-Resolution Face Recognition of Multi-Scale Blocking CS-LBP and Weighted PCA. International Journal of Pattern Recognition and Artificial Intelligence, 30(08), pp. 1–13 (2016)
13. Bodla, N., Zheng, J., Xu, H., Chen, J.C., Castillo, C., Chellappa, R.: Deep Heterogeneous Feature Fusion for Template-Based Face Recognition. In: IEEE Winter Conference on Applications of Computer Vision, pp. 586–595 (2017).
14. de Freitas Pereira, T., Marcel, S.: Heterogeneous Face Recognition using Inter-Session Variability Modelling. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 111-118 (2016).
15. Xing, X., Wang, K.: Couple manifold discriminant analysis with bipartite graph embedding for low-resolution face recognition. Signal Processing, 125, pp.329–335 (2016).
16. Wang, Z., Chang, S., Yang, Y., Liu, D., Huang, T.S.: Studying very low resolution recognition using deep networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 4792–4800 (2016).
17. Reale, C., Nasrabadi, N.M., Kwon, H.,Chellappa, R.: Seeing the Forest from the Trees: A Holistic Approach to Near-infrared Heterogeneous Face Recognition. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 54–62 (2016).
18. He, R., Wu, X., Sun, Z., Tan, T.: Learning Invariant Deep Representation for NIR-VIS Face Recognition. In: Association for Advancements in Artificial Intelligence, pp. 2000–2006 (2017).
19. Sarfraz, M.S., Stiefelhagen, R.: Deep perceptual mapping for thermal to visible face recognition. arXiv preprint arXiv:1507.02879 (2015).
20. Liu, X., Song, L., Wu, X., Tan, T.: Transferring deep representation for NIR-VIS heterogeneous face recognition. In: International Conference on Biometrics ,pp. 1–8 (2016).
21. Saxena, S., Verbeek, J.: Heterogeneous face recognition with cnns. In: European Conference on Computer Vision Workshops, pp. 483–491. (2016).
22. Wang, S., Zhang, L., Liang, Y., Pan, Q.: Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2216–2223 (2012).
23. Huang, D.A., Frank Wang, Y.C.: Coupled dictionary and feature space learning with applications to cross-domain image synthesis and recognition. In: IEEE International Conference on Computer Vision, pp. 2496–2503 (2013).

24. Bao, C., Cai, J.F., Ji, H.: Fast sparsity-based orthogonal dictionary learning for image restoration. In: IEEE International Conference on Computer Vision, pp. 3384–3391 (2013).
25. Fernando, B., Habrard, A., Sebban, M., Tuytelaars, T.: Unsupervised visual domain adaptation using subspace alignment. In: IEEE International Conference on Computer Vision, pp. 2960–2967 (2013).
26. Koestinger, M., Hirzer, M., Wohlhart, P., Roth, P.M., Bischof, H.: Large scale metric learning from equivalence constraints. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 2288–2295 (2012).
27. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep Face Recognition. In: British Machine Vision Conference, (2015).
28. Lu, J., Liong, V.E., Zhou, X., Zhou, J.: Learning compact binary face descriptor for face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(10), pp. 2041-2056 (2015).
29. Sharma, A., Kumar, A., Daume, H., Jacobs, D.W.: Generalized multiview analysis: A discriminative latent space. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2160–2167 (2012).
30. Rasiwasia, N., Mahajan, D., Mahadevan, V., Aggarwal, G.: Cluster canonical correlation analysis. In Artificial Intelligence and Statistics, pp. 823–831 (2014).
31. Kan, M., Shan, S., Zhang, H., Lao, S., Chen, X.: Multi-view discriminant analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(1), pp.188–194 (2016).
32. Lopez-Paz, D., Sra, S., Smola, A., Ghahramani, Z., Schlkopf, B.: Randomized nonlinear component analysis. In: International Conference on Machine Learning, pp. 1359–1367 (2014).
33. Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(6), pp. 1127–1133 (2010).
34. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 19(11), pp. 2861–2873 (2010).
35. Huang, J.B., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 5197–5206 (2015).